MNRAS 000, 1–7 (2015) Preprint 8 November 2022 Compiled using MNRAS LATEX style file v3.0

A Bayesian Approach to RFI Mitigation

S.A.K. Leeney,[1]★ W.J. Handley,[2] E. de Lera Acedo[2,3]
[1]University of Cambridge, Cavendish Radio Cosmology, J. J. Thomson Avenue, Cambridge, CB3 0HE, UK
[2]Department, Institution, Street Address, City Postal Code, Country
[3]Another Department, Different Institution, Street Address, City Postal Code, Country

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

In this paper we propose a first of its kind RFI mitigation algorithm that takes a Bayesian Approach, where contaminated data is both flagged and managed as part of a single step fitting process. We provide an analytical proof of our methods, which are also tested and shown to be effective on both a simple toy model and when incorporated into the Bayesian data analysis pipeline for REACH, a modern radio telescope.

Key words: keyword1 – keyword2 – keyword3

1 INTRODUCTION

The field of radio astronomy is growing rapidly. Since the year 2000, the number of known radio sources has grown from a few hundred thousand to over two million. This is expected to increase further to 70 million with the development of the ASKAP-EMU Johnston et al. (2008) telescope, and beyond with the development of the SKA Bourke et al. (2015). Furthermore, global 21cm experiments such as EDGES Bowman et al. (2018), REACH de Lera Acedo et al. (2022) and SARAS Singh et al. (2017) aim to probe into yet unseen redshift ranges to gain knowledge on the Cosmic Dawn, Dark Ages and Re-ionisation Era.

A major constraint on the design of all such telescopes is Radio Frequency Interference (RFI), which is typically orders of magnitude brighter in amplitude than the sky signal and cannot be modelled as Gaussian noise. RFI is typically anthropogenic, narrow band and comes in various species that can be either constant in time or transient Ellingson & Lewis (2006). Sources of RFI include FM Radio and Digital TV signals.

Building antennae in remote locations on earth can alleviate but never completely avoid such nuisance signals. Furthermore, RFI always increases the amplitude of the detected signal and is often transient in nature. Therefore, simply averaging the sky over long time spans (in global experiments, for example) is not sufficient to prevent contaminated data leading to systematic errors.

Currently, RFI is flagged and then excised prior to and separately from the data analysis stage of of the radio experiment. In older telescopes this may have been done manually by the astronomer. However with modern telescopes gathering terabytes of data per second, an automated approach is required. Various approaches have been taken in the past including wavelet based methods Oslick et al. (1998), CUMSUM Baan et al. (2004), and the AOFLAGGER Offringa (2010) package used for LOFAR, which implements Singular Value Decomposition Offringa et al. (2010). Another approach, used by the HERA project, is to flag the data using a watershed segmentation algorithm Kerrigan et al. (2019). More recently, Deep Learning models using convolutional neural networks trained on manually flagged and/or and simulated data Akeret et al. (2017); Vafaei Sadr et al. (2020); Sun et al. (2022) have been shown to be able to effectively flag varying species of RFI.

In this paper we propose a first of its kind RFI mitigation algorithm that takes a Bayesian Approach, where RFI is both flagged and managed within probability space. The paper is structured as follows: we introduce the basic theory behind Bayesian Inference and Nested Sampling section 2.1, then we introduce our Bayesian RFI Correction method section 2.2 followed by an analysis of said model when tested on a simple toy model section 3. The model is then tested in a real use case in the REACH pipeline section 4 and finally we conclude in section 5.

2 THEORY

2.1 Bayesian Inference

Bayesian methods can be used to perform parameter estimation and model comparison. A model $\mathcal{M}$ uses data $\mathcal{D}$ to infer its free parameters θ. Using Bayes Theorem,

$$P(\mathcal{D}|\theta) \times P(\theta) = P(\theta|\mathcal{D}) \times P(\mathcal{D}), \quad (1)$$

$$\mathcal{L} \times \pi = \mathcal{P} \times \mathcal{Z}, \quad (2)$$

the prior π is updated onto the posterior $\mathcal{P}$ in light of the likelihood $\mathcal{L}$ and furthermore the Bayesian Evidence $\mathcal{Z}$ can be inferred by computing the integral

$$\mathcal{Z} = \int \mathcal{L}(\theta) \times \pi(\theta)\, d\theta. \quad (3)$$

In practice, $\mathcal{P}$ and $\mathcal{Z}$ are determined simultaneously using the Nested Sampling algorithm POLYCHORD Handley et al. (2015). A series of points generated within π are updated such that they sequentially contract around the peak(s) of the likelihood, forming the posterior which can be used to generate parameter estimations. The artifacts of this process can then be used to compute $\mathcal{Z}$, which is used for model

★ E-mail: sakl2@cam.ac.uk (KTS)

comparison. For a more detailed description of Bayesian Inference and Nested Sampling see Sivia & Skilling (2006).

2.2 RFI correction likelihood

A data point contaminated by RFI can be considered corrupted. Any information relevant to the model is lost and furthermore it cannot be modeled as Gaussian noise. Assuming $\mathcal{D}$ is uncorrelated, the likelihood

$$\mathcal{L} = P(\mathcal{D}|\theta) = \prod_i \mathcal{L}_i(\theta), = \prod_i P(\mathcal{D}_i|\theta) \tag{4}$$

where i represents the i'th data point, is insufficient to model such contaminated data. It is therefore necessary to model the likelihood that each data point is corrupted.

We therefore introduce a piecewise likelihood including the possibility of corruption of data

$$P(\mathcal{D}_i|\theta) = \begin{cases} \mathcal{L}_i(\theta) & : \text{uncorrupted} \\ \Delta^{-1}[0 < \mathcal{D}_i < \Delta] & : \text{corrupted} \end{cases} \tag{5}$$

where we model corruption as the data becoming completely unreliable, and therefore being distributed uniformly within some range Δ. An efficient way to write this likelihood is

$$P(\mathcal{D}|\theta, \varepsilon) = \prod_i \mathcal{L}_i^{\varepsilon_i} \Delta^{\varepsilon_i - 1} \tag{6}$$

where the Boolean mask vector ε has a ith component which takes the value 1 if the datum i is uncorrupted and value 0 if corrupted.

We do not know before the data arrive whether or not they are corrupted by RFI, but we may infer it in a Bayesian fashion, by ascribing a Bernoulli prior probability p_i of corruption i.e:

$$P(\varepsilon_i) = p_i^{(1-\varepsilon_i)}(1 - p_i)^{\varepsilon_i}. \tag{7}$$

It should be noted that above we assume the a-priori probability that each bin contains RFI is uncorrelated, i.e $P(\varepsilon) = \prod_i P(\varepsilon_i)$, which in practice this will almost certainly not be true. We will discuss later the extent to which this assumption can be considered valid.

Multiplying eqs. (6) and (7) yields

$$P(\mathcal{D}, \varepsilon|\theta) = \prod_i [\mathcal{L}_i(1 - p_i)]^{\varepsilon_i} [p_i/\Delta]^{(1-\varepsilon_i)} \tag{8}$$

and to recover a likelihood independent of ε we formally can marginalise out:

$$P(\mathcal{D}|\theta) = \sum_{\varepsilon \in \{0,1\}^N} P(\mathcal{D}, \varepsilon|\theta) \tag{9}$$

$$= \sum_{\varepsilon \in \{0,1\}^N} \prod_i [\mathcal{L}_i(1 - p_i)]^{\varepsilon_i} [p_i/\Delta_i]^{(1-\varepsilon_i)}. \tag{10}$$

This would require the computation of the all 2^N terms in eq. (10). For realistic values of N, this computation becomes impractical. However, if it is assumed that the most likely model (i.e the maximum term in eq. (10)) dominates over the next to leading order terms, we can make the approximation

$$P(\mathcal{D}, \varepsilon|\theta) \approx \delta_{\varepsilon \varepsilon^{\max}} \times P(\mathcal{D}, \varepsilon^{\max}|\theta) \tag{11}$$

where δ_{ij} is the usual Kroneker delta function, and $\varepsilon^{\max}$ is the mask vector which maximises the likelihood $P(D, \varepsilon|\theta)$, namely:

$$\varepsilon_i^{\max} = \begin{cases} 1, & \mathcal{L}_i(1 - p_i) > p_i/\Delta_i \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

Under this approximation we find that the sum in eq. (10) becomes

$$P(\mathcal{D}|\theta) \approx P(\mathcal{D}, \varepsilon^{\max}|\theta). \tag{13}$$

In practice the approximation in eq. (13) is only valid if the next to leading order term is much smaller, such that

$$P(\mathcal{D}|\theta, \varepsilon_{\max}) \gg \max_j P(\mathcal{D}|\theta, \varepsilon^{(j)}), \tag{14}$$

where $\varepsilon^{(j)}$ is $\varepsilon^{\max}$ with its jth bit flipped:

$$\varepsilon_k^{(j)} = \begin{cases} 1 - \varepsilon_k^{\max} & k = j \\ \varepsilon_k^{\max} & k \neq j \end{cases} \tag{15}$$

and we may use eq. (14) as a consistency check.

To summarise, we can correct for RFI under these approximations by replacing the original likelihood $\mathcal{L} = \prod_i \mathcal{L}_i$ in eq. (4) with

$$P(\mathcal{D}|\theta) = \prod_i [\mathcal{L}_i(1 - p_i)]^{\varepsilon_i^{\max}} [p_i/\Delta]^{(1-\varepsilon_i^{\max})} \tag{16}$$

where $\varepsilon^{\max}$ is defined by eq. (12).

2.3 Computing the posterior

The posterior and evidence are computed from eq. (16) via Nested Sampling (although any numerical Bayesian sampling method could be used). Taking logs for convenience gives

$$\log P(\mathcal{D}|\theta) = \sum_i ([\log \mathcal{L}_i + \log(1 - p_i)]\varepsilon^{\max} + [\log p_i - \log \Delta](1 - \varepsilon_i^{\max})), \tag{17}$$

yielding a masked chi squared like term which can be used to distinguish whether there is a statistically significant difference between the classes of data, i.e corrupted or non corrupted. Furthermore the second term in eq. (17) introduces an Occam's penalty. Each time a data point is predicted to be RFI it is replaced with said penalty rather than being completely removed. Without this term, the likelihood where all data points are flagged would be more simple and thus 'more likely' than all other possibilities.

We compute this by imposing the condition in eq. (12) on eq. (17) as follows,

$$\log P(\mathcal{D}|\theta) = \begin{cases} \log \mathcal{L}_i + \log(1 - p_i), & [\log \mathcal{L}_i + \log(1 - p_i) \\ & > \log p_i - \log \Delta] \\ \log p_i - \log \Delta, & \text{otherwise.} \end{cases} \tag{18}$$

The effect that this correction has on a log-likelihood with a corrupted datapoint is shown in fig. 1, where the likelihood of the corrupted point is replaced, reducing the impact that this point has on the subsequent posterior. A penalty is also imposed here, rather than deleting the datapoint altogether, moving the point onto the grey dashed line so it is still 'less likely' than all other points. Without this penaly, flagging all points would lead to a loglikelihood of 0. Therefore flagging all points would always be preferable, even compared to flagging the correct points.

The corrected likelihood is then updated iteratively via the selected Bayesian sampling method, compressing the prior onto the posterior while simultaneously correcting for RFI. One may also notice that the right hand side of eq. (17) is a Logit function. Logit functions are used routinely as an activation function in binary classification tasks, which hints at the potential of a future extension of this work using machine learning.

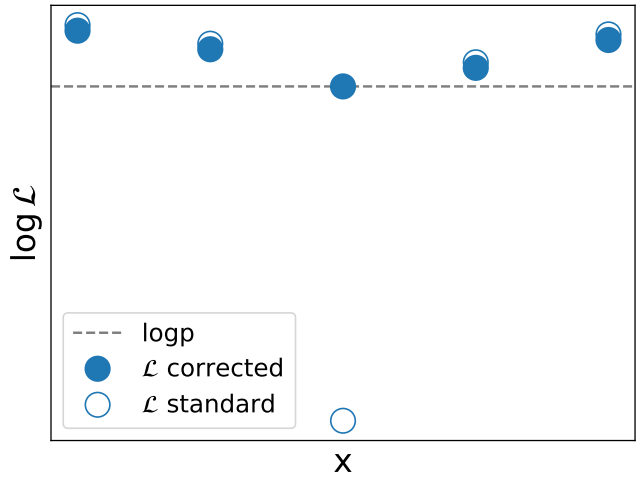

Figure 1. Showing the effect of the correction on the likelihood for a simple 5 point data set where the 3rd point is corrupted.

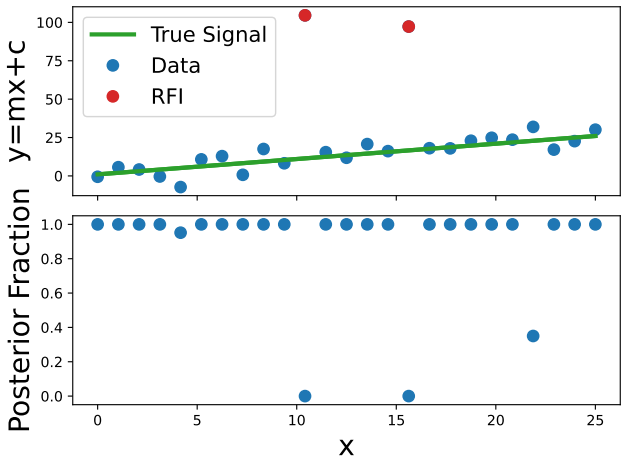

Figure 2. The top pane shows the dataset, with the true signal marked in green and points corrupted by RFI marked in red. The bottom pane shows the mean posterior on ε, produced during the Nested Sampling run.

3 TOY EXAMPLE

We will initially test this approach on a simple toy model, consisting of a straight line with Gaussian noise and RFI injected. Then, we move onto a more realistic and complex case in section 4.

3.1 Dataset

Two datasets of 25 data points are generated for comparison. The first is a line with Gaussian noise $\sigma = 5$, gradient $m = 1$ and intercept $c = 1$. It does not contain RFI thus can be used as a ground truth. The second dataset is the same, with the random noise constrained so the data is reproducible, but with two RFI spikes injected. This is shown in the top pane of fig. 2.

3.2 Initial Testing

We fit the data described in section 3.1 in a Bayesian sense and attempt to recover the two free parameters m and c using the correcting likelihood in eq. (18) with

$$\mathcal{L}_i = -\frac{\log(2\pi\sigma^2)}{2} - \frac{[y_i - y_{\mathcal{S}}(x_i)]^2}{2\sigma^2}. \tag{19}$$

Here σ is the Gaussian noise, y_i is the simulated data and $y_{\mathcal{S}}(x_i)$ is the output when the estimated parameters from the current sampling iteration are used to compute the model $y_i = mx_i + c$. Sampling the posterior $P(\theta|\mathcal{D})$ via a Bayesian sampling procedure, the prior distribution is compressed and the belief that each point fits (or is corrupted and does not fit) the model is updated. Thus by evaluating the subsequent posterior on ε, as shown in fig. 2, we can assess how many times across the entire sampling run each data point was believed to fit (non corrupted) or not fit (corrupted) the model. As seen in fig. 2, it is evident that the points containing RFI make up a near zero fraction of the posterior, because as the prior was iteratively updated onto the posterior they were frequently believed to be corrupted. Conversely, points that do not contain RFI often fit the model and as such contribute significantly to the final posterior distribution. There are also some points that lie somewhere in between, which the model is less confident are uncontaminated. When comparing the actual data in the top pane to the probability space view in the bottom pane, it appears as though the points the model has less confidence in are the ones that deviate the most from the true signal due to Gaussian noise.

It should be emphasised that although ε_i is constrained to binary values, the subsequent mask on ε is not. Unlike traditional RFI flagging algorithms, points are not simply flagged as 'RFI' and 'Not RFI'. The mask takes the weighted mean across the posterior. Thus, points more likely to contain RFI will have less 'impact' on the final posterior distribution than points believed to be uncontaminated. In practice, as seen in fig. 2, points predicted to contain RFI will have a non-zero contribution on the subsequent posterior. However, it is possible that points which the model is less confident in, such as $x = 19$ and $x = 9$ will have lesser but non zero contribution to the final distribution. The mask could be thought of as being slightly opaque to these data points, accounting for the models uncertainty. The ability to incorporate the models confidence in its correction directly into the subsequent parameter estimations and model comparison makes this mitigation approach unique in comparison with its counterparts.

3.3 Model Evaluation

To further examine the effectiveness of the correction it is necessary to develop a simple toy model, similar to the above but also simulating various other scenarios. Both datasets described in section 3.1 will be evaluated when fit using the likelihood capable of correcting for RFI. They will also be fit using a traditional likelihood, which cannot account for RFI. This will generate a total of four posterior distributions for comparison.

We initially fit a baseline case to be used as a ground truth, using the clean dataset and a traditional likelihood that does not facilitate the modeling of RFI. The new likelihood, capable of correcting for RFI is then used to fit the clean dataset, to confirm that it performs correctly even in the absence of RFI. The dataset containing RFI is then fit using the new likelihood and the traditional likelihood, to examine how much better the model performs when the correction is applied. All but the contaminated, uncorrected case would be expected to perform similarly if RFI has been effectively mitigated. It should be noted however that from a Bayesian standpoint, the simplest model will always be preferable. So, for the clean dataset

it would be expected that the standard likelihood would be preferred slightly over the correcting likelihood.

Fig. 3 shows the parameter distributions inferred from the data in section 3.1 in the four cases described above and their associated 1σ and 2σ confidences. The bottom left pane shows that both m and c are inferred to within 1σ of their true values. The 'No RFI, No Correction' case is similar to the 'RFI, Corrected' case, indicating that the model has effectively corrected the corrupted regions of data. This is particularly evident when comparing these two cases to the uncorrected case.

According the Bayes theorem, the simplest model will always be preferred. This is evidenced when comparing the Bayes Factor computed from the log evidences in table section 3.3. For the uncontaminated data, the model that does not correct for RFI is slightly preferred. Conversely the correcting model is strongly preferred on the contaminated data indicating that the correction is working effectively.

Table 1	No RFI	RFI
No Correction ($\log \mathcal{Z}$)	-79	-474
Correction ($\log \mathcal{Z}$)	-80	-86
log Bayes Factor	1	-388

It is also useful to view the actual functions plotted from the inferred parameter estimations and the associated confidences. As seen in fig. 4, when RFI is not corrected, the true value is well outside the 1σ and sometimes 2σ confidence bounds. Conversely the other three cases fit almost entirely within the 1σ bounds, indicating that the RFI has been mitigated.

3.4 Evaluating the p dependence

Proper selection of the probability thresholding term $\log p$ is essential to the efficacy of the mitigation process. Set too small, the model will require such a high confidence of a data point fitting the model that deviations due to Gaussian noise will be predicted as RFI. Set to low, nothing will be corrected. From a Bayesian standpoint $\log p$ would be set to represent our prior degree of belief in their being RFI in each datum.

We firstly assess the question in a qualitative manor, viewing the posterior on $\varepsilon_{\max}$ for a selection of $\log p$ values. This is computed by evaluating $P(\mathcal{D}|\theta, \varepsilon_{\max})$, avoiding the sum to preserve each data point in x for each posterior sample. We then weighted average across the samples, giving the probability that each data point fits the model. More specifically, the probability that each data point is uncontaminated by RFI.

When $\log p$ is set appropriately as seen in the left pane of fig. 5, the model has a high confidence in its predictions. Across the sampling run (here, we used Nested Sampling) the probability threshold frequently laid between the likelihood for contaminated and uncontaminated datapoints, while not being so high (and therefore sensitive) as to flag any deviations caused by higher order Gaussian noise. At $\log p = -3.0$ the RFI still appears to have been accurately modeled, however the confidence in some of the predictions are lower. As $\log p$ increases further to -1.9, the model breaks down. Points are predicted incorrectly in many of the sampling iterations and as such the model is less able to confidently predict their class. Viewing fig. 5 from left to right, we observe the effects of this thresholding term moving to the peak of the likelihood, sweeping up points the model is less confident in due to Gaussian noise. We assess the p dependence while varying $\log p$ as a function of the RMSE on the fit generated from the parameter estimations, the log Bayesian Evidence and the mean number of points flagged across all samples.

For $\log p$ close to zero, the RMSE is high and we observe in fig. 6 that the model has been unable to generate accurate parameter estimations. Here the threshold is so high that the model is more confident that any of the points are RFI than non RFI. This matches the corresponding low evidence. The RMSE drops sharply as $\log p$ decreases below -0.1 to near its minimum. Here, the model still incorrectly flags ≈ 5 data points, showing that the model is able to generate accurate parameter estimations while over flagging, indicating it is insensitive to false positives. As $\log p$ decreases further, the model is more able to distinguish between higher order Gaussian noise and as such the average number of points predicted to be RFI approaches the true value. As this happens the evidence also reaches its maximum, which indicates that the Bayesian Evidence is appropriately showing how well each of the many models created by different $\log p$ values fit the data.

3.5 To what extent is $P(\mathcal{D}|\theta) \approx P(\mathcal{D}|\theta, \varepsilon_{\max})$ valid?

A key assumption is made in eq. (13) is that the leading order term, eq. (17), has a magnitude considerably larger than all the other possible models for $\varepsilon \in (0, 1)^N$. If this approximation is not valid, the mathematical framework supporting the implementation behind this approach is invalid.

Therefore it is necessary to test the validity of this approximation by computing eq. (18) and comparing the result with the next leading order term as calculated by eq. (14). For $-5 < \log p < -0.01$, the results are displayed in the bottom pane of fig. 6. The difference between the maximised term and the next leading order term is 9 exponential log units at peak $\log \mathcal{Z}$ and increases linearly beyond this. As such, depending on the $\log p$ selection strategy, the optimal mask is 9 orders of magnitude more likely than the next most likely mask.

3.6 Selection Strategy for $\log p$

Various selection strategies could be taken to select the optimal $\log p$ value. For each $\log p$, the model changes. As such, selecting the $\log p$ that maximises the evidence seems to be the most obvious selection strategy. In the case of the toy model, the peak $\log \mathcal{Z}$ occurs where $\log p = -3.2$ as shown in fig. 6.

Another possible strategy could be to select $\log p$ where the number of points flagged is at its minimum. One could also make the selection where the difference between the maximum term in eq. (7) and the next leading order term is judged to be suitably large to make the approximation in eq. (13) valid. We recommend using a combination of this, and the Bayesian Evidence. For example, 'where the approximation in eq. (13) is valid, select $\log p$ such that $\log \mathcal{Z}$ is maximised'. It is also possible to ascribe a prior to $\log p$, fully automating the approach. This will be examined further in future works.

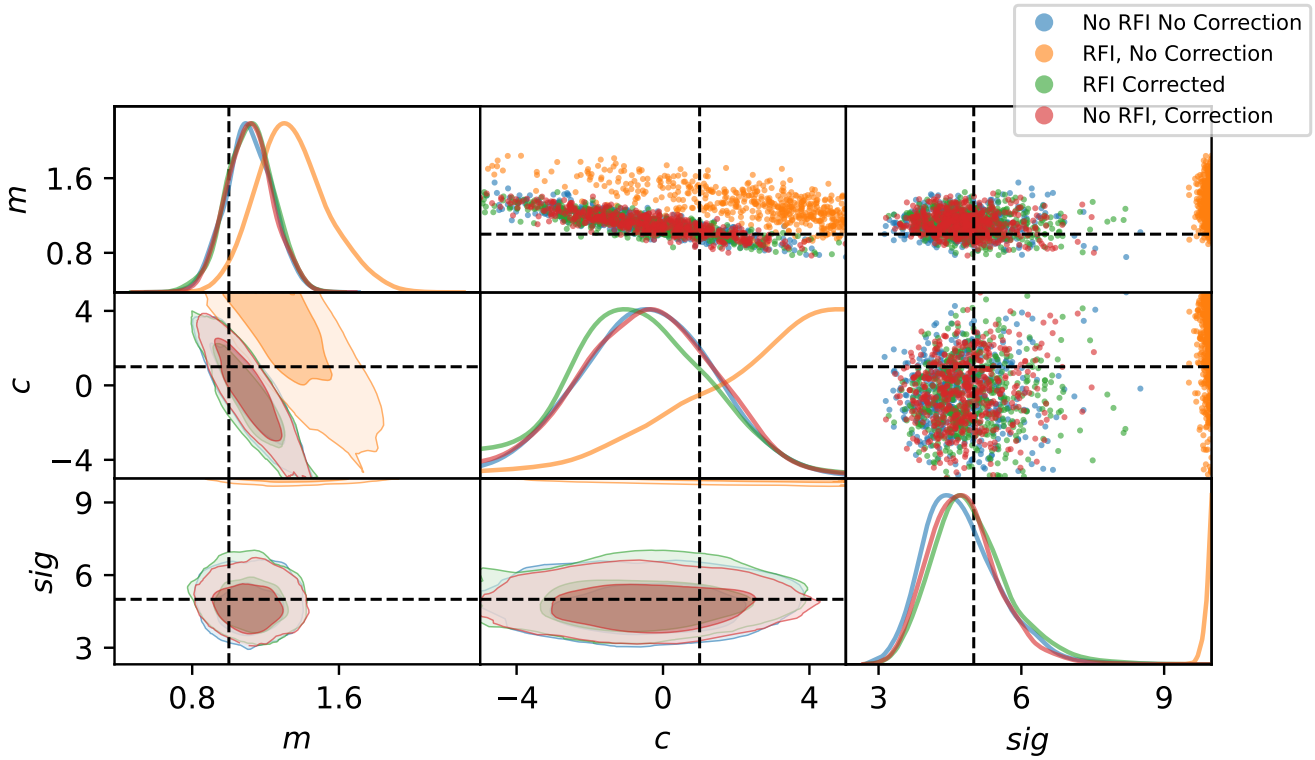

Figure 3. Showing the parameter distributions inferred from the dataset described in Section 3.1. The top left to bottom right panes show probability distribution functions for m, c and σ, respectively. Plots generated using posterior plotting tool ANESTHETIC Handley (2019).

4 REACH EXAMPLE

Finally, we examine a real use case for this method. The REACH de Lera Acedo et al. (2022) radio telescope is designed to detect the faint 21cm signal from the cosmic dawn. This signal is 5 orders of magnitude dimmer than the foreground, therefore a highly precise measurement is required. The REACH data analysis pipeline takes a Bayesian Approach to antenna calibration, foreground modelling and chromaticity correction Anstey et al. (2021) and as such is a useful environment to test our RFI mitigation framework. The 21cm signal is expected to take the shape of an inverted Gaussian, so the model takes the form

$$f(x) = A \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{20}$$

with center frequency μ, standard deviation σ and magnitude A all free parameters. The four cases discussed in section 3 are then examined, but this time on a simulated sky data set containing a 21cm signal with two RFI spikes injected. The evidences and subsequent log Bayes Factors are shown in Table 2 below.

Table 2	No RFI	RFI
No Correction ($\log \mathcal{Z}$)	294	-976661
Correction ($\log \mathcal{Z}$)	293	269
Bayes Factor	1	-976930

The No RFI, Correction and ground truth cases are very similar with the simpler (ground truth) case marginally preferred as expected. The RFI, Corrected case is again similar with a slightly lower evidence due to the penalties incurred during the corrections. The above is also evident viewing fig. 7. The reconstructed signal is within 2σ of the true signal for all but the RFI, No Correction case.

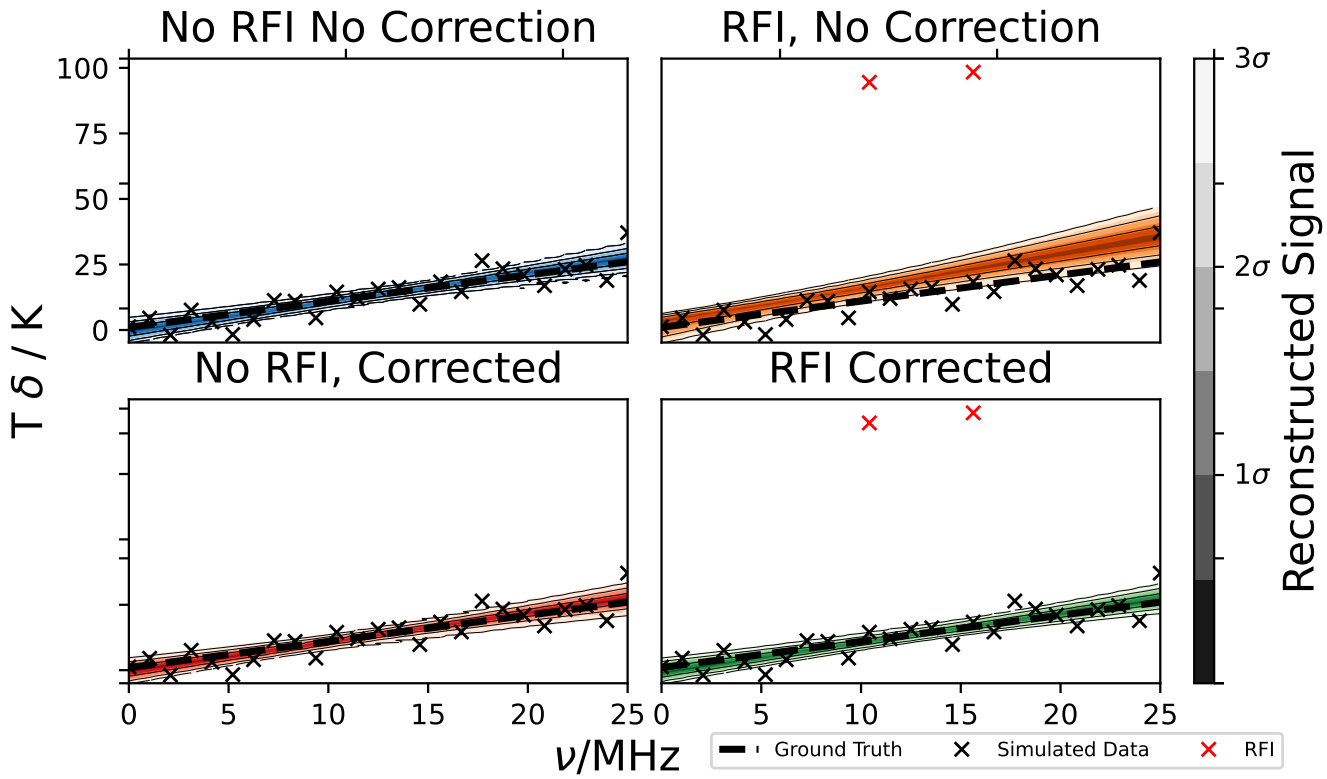

Figure 4. Showing the inferred parameter estimations in a contour plot, where darker tones indicate higher σ confidence in the parameter estimation. Generated from the dataset described in Section 3.1. The plots are generated using the functional posterior plotter fgivenx Handley (2018).

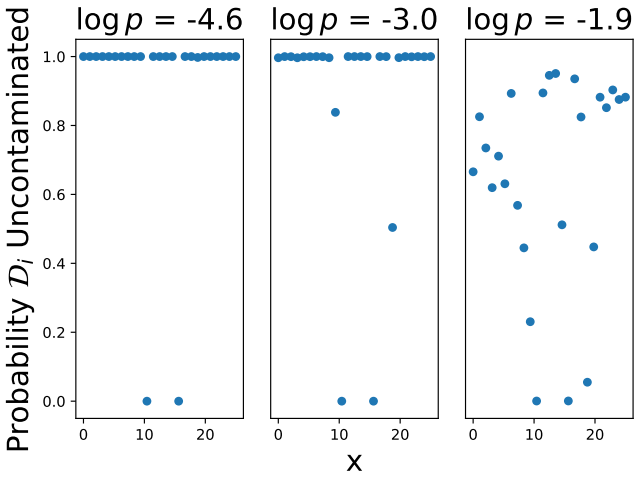

Figure 5. Showing the mean mask where log p is varied from its optimal value, to the point at which the model breaks down.

5 CONCLUSIONS

In this paper, which serves as a proof of concept, we show that that RFI can be both flagged and corrected in a truly Bayesian Sense within probability space. This first of its kind approach to RFI mitigation may be a highly effective tool for future radio telescopes, given the growing prevalence of modern data analysis pipelines taking a Bayesian Approach and the need for a system that can be incorporated into these increasingly complex data analysis systems. Our methods can effectively perform inference on data hampered by RFI, as shown when tested on a simple toy model and furthermore in a real use case in the pipeline of a global 21cm experiment, all as part of a single step Bayesian fitting process.

In future works, it will be necessary to further examine to approximation made in eq. (13) to confirm that it is valid in other more complex scenario's. Further analysis is also required to assess the efficacy of this technique when compared to other modern RFI mitigation approaches. As mentioned in section 2.2 we assume that each data point is separable and uncorrelated. It is unlikely that this is true for real data as confirmed in Wilensky (2021), so the case where data points are dependent on each other should be investigated in future works. Development is required to incorporate this method when applied to time integrated data and furthermore on different species of RFI, such as transient and broadband.

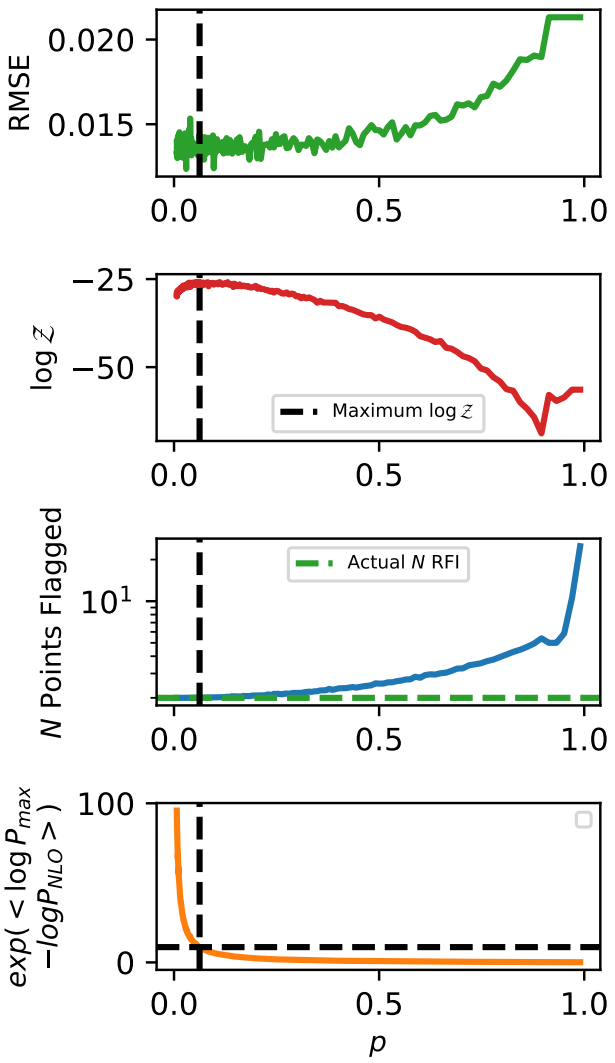

Figure 6. Shows how various methods of model evaluation change as the thresholding term is varied. From top to bottom: the RMSE, then the log of the Bayesian Evidence, then the weighted mean number of points flagged and finally the difference between the log-likelihood for $\varepsilon_{\max}$ and its next leading order term.

REFERENCES

Akeret J., Chang C., Lucchi A., Refregier A., 2017, Astronomy and computing, 18, 35

Anstey D., de Lera Acedo E., Handley W., 2021, Monthly Notices of the Royal Astronomical Society, 506, 2041

Baan W., Fridman P., Millenaar R., 2004, The Astronomical Journal, 128, 933

Bourke T., et al., 2015

Bowman J. D., Rogers A. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Nature, 555, 67

Ellingson S., Lewis B., 2006, An SKA Engineering Overview, p. 116

Handley W., 2018, The Journal of Open Source Software, 3

Handley W., 2019, The Journal of Open Source Software, 4, 1414

Handley W., Hobson M., Lasenby A., 2015, Monthly Notices of the Royal Astronomical Society, 453, 4384

Johnston S., et al., 2008, Experimental astronomy, 22, 151

Kerrigan J., et al., 2019, Monthly Notices of the Royal Astronomical Society, 488, 2605

Offringa A., 2010, Astrophysics Source Code Library, pp ascl–1010

Offringa A., De Bruyn A., Biehl M., Zaroubi S., Bernardi G., Pandey V., 2010, Monthly Notices of the Royal Astronomical Society, 405, 155

Oslick M., Linscott I. R., Maslakovic S., Twicken J. D., 1998, in Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181). pp 1537–1540

Singh S., et al., 2017, The Astrophysical Journal Letters, 845, L12

Sivia D., Skilling J., 2006, Data analysis: a Bayesian tutorial. OUP Oxford

Sun H., Deng H., Wang F., Mei Y., Xu T., Smirnov O., Deng L., Wei S., 2022, Monthly Notices of the Royal Astronomical Society, 512, 2025

Vafaei Sadr A., Bassett B. A., Oozeer N., Fantaye Y., Finlay C., 2020, Monthly Notices of the Royal Astronomical Society, 499, 379

Wilensky M., 2021, PhD thesis, University of Washington

de Lera Acedo E., et al., 2022, Nature Astronomy, pp 1–15

This paper has been typeset from a TEX/LATEX file prepared by the author.

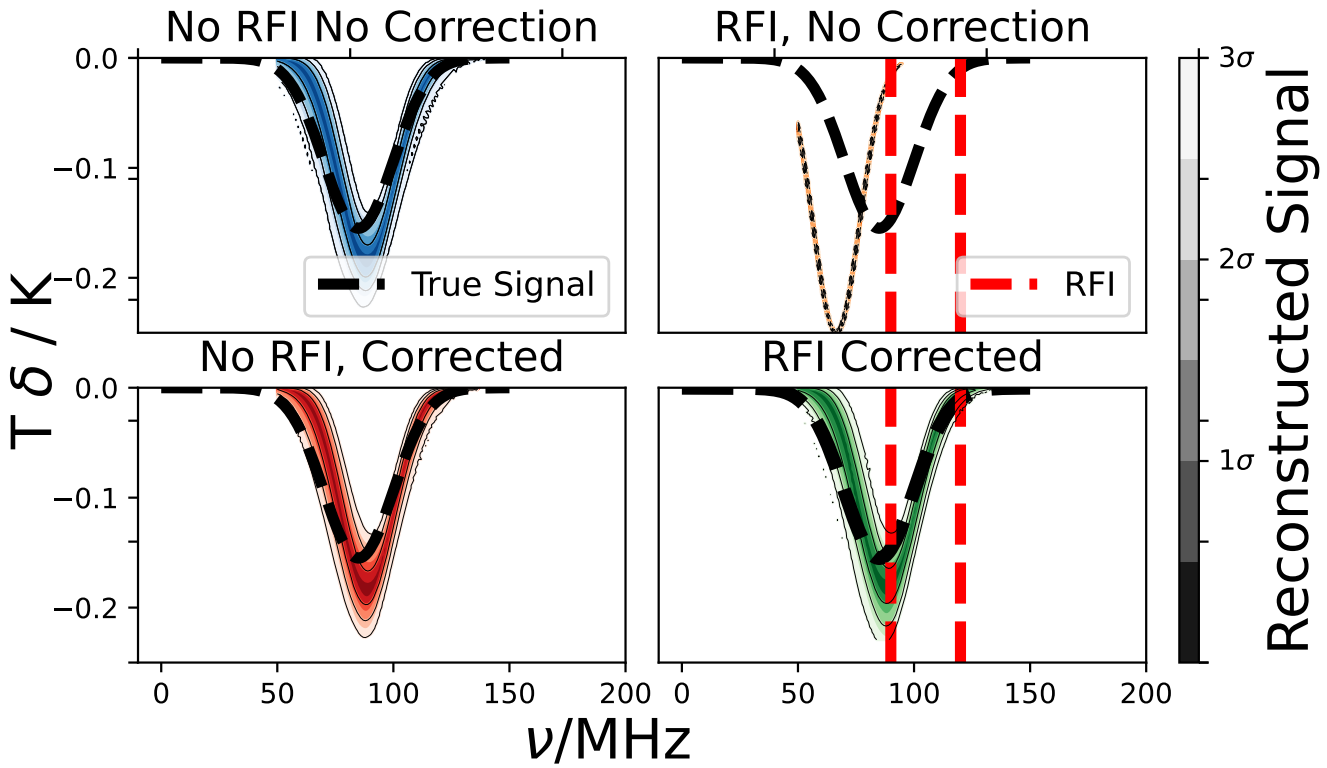

Figure 7. Showing the results when the RFI correction is applied on simulated data in the REACH data analysis pipeline. The left pane shows the case where the correction is applied to the contaminated data and the right pane shows the contaminated data fit with the standard likelihood.